

# STATISTIQUE MATHÉMATIQUE – EXAMEN

M1 Mathématiques Fondamentales, Université Paris-Saclay  
2025-2026

Lundi 13 avril 2026  
14h – 17h

*Une feuille A4 recto manuscrite est autorisée en tant que support. La calculatrice est autorisée.*

## *Avant de commencer :*

- *Le sujet comporte deux problèmes. Il est demandé de les traiter dans l'ordre sur la copie.*
- *Une attention particulière sera portée aux questions d'interprétation, à la rigueur et à la précision de la rédaction.*
- *Le sujet est long et n'est pas pensé pour être terminé dans son intégralité dans le temps imparti. Des résultats pourront être admis d'une question à l'autre, à la condition de l'écrire clairement.*

*Bon courage et amusez-vous !*

Dans les deux problèmes,  $\|\cdot\|$  désigne la norme euclidienne canonique de l'espace  $\mathbb{R}^d$ , avec  $d$  clair dans le contexte.

### Problème 1 – Modèle autorégressif gaussien d'ordre $p$

On admet dans tout ce problème le résultat classique suivant (décomposition de Cholesky). Si  $A$  est une matrice symétrique définie positive, il existe une matrice réelle triangulaire inférieure  $L$  telle que  $A = LL^\top$ . De plus, cette décomposition est unique si l'on impose à  $L$  d'avoir des coefficients diagonaux strictement positifs.

On se place dans le modèle où l'on considère une série temporelle réelle aléatoire  $X_0, X_1, \dots, X_T$  où  $T \geq 1$ . Ce modèle<sup>1</sup> est défini récursivement de la façon suivante : pour tout  $t \geq 1$ ,

$$X_t = \beta_0^* + \beta_1^* X_{t-1} + \dots + \beta_p^* X_{t-p} + \varepsilon_t,$$

avec  $(\varepsilon_t)_{1 \leq t \leq T}$  des bruits i.i.d. de loi  $\mathcal{N}(0, \sigma^2)$ , et  $\beta^* = (\beta_0^*, \dots, \beta_p^*)^\top \in \mathbb{R}^{p+1}$  un vecteur de paramètres. Par convention on prendra  $X_m = 0$  si  $m \leq 0$  dans la définition ci-dessus. On notera  $X = (X_1, \dots, X_T)^\top \in \mathbb{R}^T$ ,  $\varepsilon = (\varepsilon_1, \dots, \varepsilon_T)^\top \in \mathbb{R}^T$ . Les paramètres inconnus du modèle sont  $\beta^*$  et  $\sigma^2$ .

1. On note

$$M^* = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -\beta_1^* & 1 & 0 & \cdots & 0 \\ -\beta_2^* & -\beta_1^* & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\beta_p^* & -\beta_{p-1}^* & -\beta_{p-2}^* & \cdots & 1 \end{pmatrix}$$

1.a. On suppose  $T = p + 1$ . Montrer que  $M^* X = \beta_0^* \mathbf{1}_{p+1} + \varepsilon$ , avec  $\mathbf{1}_{p+1}$  le vecteur de  $\mathbb{R}^{p+1}$  contenant uniquement des 1.

1.b. On suppose toujours  $T = p + 1$ . En déduire la loi de  $X$  en fonction de  $\beta^*$  et  $M^*$ .

1.c. En déduire que si  $T \geq p + 1$ , le modèle est identifiable.

2. Que dire de l'identifiabilité du modèle si  $T < p + 1$ ?

Dans la suite du problème, on suppose  $T \geq p + 1$ . On note  $\Phi(X)$  la matrice de  $\mathbb{R}^{T \times (p+1)}$  définie par

$$\Phi(X) = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & X_1 & 0 & \cdots & 0 \\ 1 & X_2 & X_1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_p & X_{p-1} & \cdots & X_1 \\ 1 & X_{p+1} & X_p & \cdots & X_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & X_{T-1} & X_{T-2} & \cdots & X_{T-p} \end{pmatrix}.$$

Avec ces notations, le modèle s'écrit matriciellement :

$$X = \Phi(X)\beta^* + \varepsilon.$$

Notons qu'ici,  $\Phi(X)$  est aussi aléatoire : le modèle écrit n'est pas donc pas un modèle linéaire au sens du cours.

<sup>1</sup>On parle de *modèle autorégressif gaussien d'ordre  $p$* . C'est un modèle classique en séries temporelles : l'état d'une mesure au temps  $t$  dépend des  $p$  dernières mesures, à bruit additif près.

3. Ecrire la vraisemblance  $L(X, \beta, \sigma^2)$  du modèle, et en déduire que tout estimateur du maximum de vraisemblance  $\hat{\beta}$  de  $\beta^*$  vérifie

$$\hat{\beta} \in \arg \min_{\beta \in \mathbb{R}^p} \|X - \Phi(X)\beta\|^2.$$

4. Montrer que sous nos hypothèses, le problème de minimisation précédent admet presque sûrement une unique solution  $\hat{\beta}_{MC}$  (dite des moindres carrés) donnée par

$$\hat{\beta}_{MC} := (\Phi(X)^\top \Phi(X))^{-1} \Phi(X)^\top X.$$

L'étude de la loi de  $\hat{\beta}_{MC}$  à  $T$  fini est complexe, à cause de la structure de dépendance dans  $X$ . Cependant, on a tout de même des renseignements sur la loi asymptotique de  $\hat{\beta}_{MC}$ , sous les deux hypothèses suivantes.

- (H1).** Il existe  $Q \in \mathbb{R}^{(p+1) \times (p+1)}$  symétrique définie positive telle que, lorsque  $T \rightarrow +\infty$ ,

$$\frac{1}{T} \Phi(X)^\top \Phi(X) \rightarrow Q, \text{ p.s.}$$

- (H2).** Lorsque  $T \rightarrow +\infty$ , on a la convergence en loi:

$$\frac{1}{\sqrt{T}} \Phi(X)^\top \varepsilon \rightarrow \mathcal{N}(0, \sigma^2 Q).$$

5. Montrer que sous **(H1)** et **(H2)**,  $\sqrt{T}(\hat{\beta}_{MC} - \beta^*)$  tend en loi vers un vecteur gaussien dont on précisera les paramètres, en fonction de  $Q$  et  $\sigma^2$ .

On considère également l'hypothèse suivante.

- (H3).** Lorsque  $T \rightarrow +\infty$ ,

$$\hat{\sigma}^2 := \frac{1}{T} \|X - \Phi(X)\hat{\beta}_{MC}\|^2 \rightarrow \sigma^2, \text{ p.s.}$$

6. Le modèle étudié ici est d'ordre  $p$ , c'est-à-dire que  $X_t$  dépend des  $p$  dernières valeurs dans le passé  $X_{t-1}, \dots, X_{t-p}$ . On veut tester :

$\mathcal{H}_0$  : le modèle est d'ordre  $p-1$     contre     $\mathcal{H}_1$  : le modèle est d'ordre  $p$ .

- 6.a. Construire une statistique de test pour ce problème basée sur  $\hat{\sigma}^2$  et  $\hat{\beta}_{MC}$ , dont on donnera la loi asymptotique sous les hypothèses **(H1)**, **(H2)** et **(H3)**.
- 6.b. En déduire un test de taille asymptotique  $\alpha \in ]0, 1[$  répondant au problème sous les hypothèses **(H1)**, **(H2)** et **(H3)**.

## Problème 2 – Un résultat d'impossibilité dans le spiked covariance model

Pour tout  $d \geq 1$ , on note  $\mathbb{S}^{d-1} = \{u \in \mathbb{R}^d : \|u\|^2 = 1\}$ , et  $O(d)$  l'ensemble des matrices orthogonales réelles de taille  $d \times d$ . On désigne par  $\mu$  l'unique<sup>2</sup> mesure de probabilité sur  $\mathbb{S}^{d-1}$  invariante par rotation orthogonale :

si  $U \sim \mu$ , alors  $OU \sim \mu$  pour tout  $O \in O(d)$ .

Soit  $n \geq 1$ ,  $\sigma^2 > 0$  et  $\lambda > 0$ . On travaille dans le modèle suivant. Sous  $\mathbb{P}_0$ , on observe  $X_1, \dots, X_n$  vecteurs i.i.d. de loi  $\mathcal{N}(0, \sigma^2 I_d)$ . Sous  $\mathbb{P}_1$ , on tire  $u \sim \mu$ , puis conditionnellement

<sup>2</sup>on admet l'existence et l'unicité de  $\mu$ .

à  $u$ , on observe  $X_1, \dots, X_n$  vecteurs i.i.d. de loi  $\mathcal{N}(0, \sigma^2 I_d + \lambda uu^\top)$ . L'objectif de ce problème est l'étude du problème de test

$$\mathcal{H}_0 : (X_1, \dots, X_n) \sim \mathbb{P}_0 \quad \text{contre} \quad \mathcal{H}_1 : (X_1, \dots, X_n) \sim \mathbb{P}_1.$$

**Première partie : propriétés diverses.**

1. En français (sans aucune formule mathématique), décrivez ce qui distingue qualitativement deux nuages de points tirés sous  $\mathbb{P}_0$  et  $\mathbb{P}_1$ . Représenter la situation avec un dessin en dimension  $d = 2$ .
2. En argumentant avec des intuitions, comment la difficulté de ce problème de test vous semble-t-elle évoluer en fonction de  $\lambda$  ? de  $\sigma^2$  ? de  $d$  ? de  $n$  ?
3. Pour  $P, Q$  deux mesures de probabilité sur le même espace  $(\Omega, \mathcal{F})$  ayant des densités  $p, q$  par rapport à une mesure commune  $\nu$ . On suppose de plus que  $q > 0$  sur  $\Omega$ . On note  $\text{TV}(P, Q)$  la distance en variation totale entre  $P$  et  $Q$ . On pose

$$\chi^2(P, Q) := \int \frac{(p(x) - q(x))^2}{q(x)} d\nu,$$

bien définie dans  $\mathbb{R}_+ \cup \{+\infty\}$ . (Re)démontrer que

$$\text{TV}(P, Q) \leq \frac{1}{2} \sqrt{\chi^2(P, Q)}.$$

4.
  - 4.a. Montrer que si  $A \in \mathbb{R}^{d \times d}$  commute avec toute matrice  $O \in O(d)$ , alors  $A$  est de la forme  $\lambda I_d$  avec  $\lambda \in \mathbb{R}$ .
  - 4.b. Soit  $u \sim \mu$ . Montrer à l'aide de 4.a. que  $\mathbb{E}[uu^\top] = \frac{1}{d} I_d$ .
  - 4.c. Dédurre de ci-dessus que si  $u, v$  sont i.i.d. selon  $\mu$ , alors  $\mathbb{E}[(u^\top v)^2] = 1/d$ . On développera le carré, et on utilisera l'identité  $\mathbb{E} \circ \text{Tr} = \text{Tr} \circ \mathbb{E}$ .

5. Soit  $u \in \mathbb{S}_{d-1}$  et  $t \in \mathbb{R}$ .

- 5.a. Montrer que

$$\det(I_d + tuu^\top) = 1 + t.$$

- 5.b. Montrer que si  $t \neq -1$ ,  $I_d + tuu^\top$  est inversible d'inverse

$$(I_d + tuu^\top)^{-1} = I_d - \frac{t}{1+t} uu^\top.$$

6. Soient  $u, v \in \mathbb{S}_{d-1}$  et  $t > 0$ . On suppose  $d \geq 3$ . Montrer que

$$\det(I_d - tuu^\top - tvv^\top) = (1-t)^2 - t^2(u^\top v)^2.$$

On pourra raisonner dans  $\text{Vect}(u, v)$ .

Dans toute la suite, on note

$$\theta := \frac{\lambda}{\sigma^2}.$$

**Seconde partie : un résultat d'impossibilité.**

Pour  $u \in \mathbb{S}_{d-1}$ , on note  $\mathbb{P}_u := \mathcal{N}(0, \sigma^2 I_d + \lambda uu^\top)^{\otimes n}$ .

7. Soit  $u \in \mathbb{S}_{d-1}$ . Montrer que le rapport de vraisemblance entre  $\mathbb{P}_u$  et  $\mathbb{P}_0$ , noté  $L_u(x)$ , s'écrit :

$$L_u(x) = (1 + \theta)^{-n/2} \exp \left( \frac{\theta}{2\sigma^2(1 + \theta)} \sum_{i=1}^n (u^\top x_i)^2 \right).$$

8. On suppose dans le reste du problème que  $d \geq 3$ . On admet le résultat suivant : Si  $X \sim \mathcal{N}(0, \sigma^2 I_d)$ , et  $A \in \mathbb{R}^{d \times d}$  symétrique non inversible,

$$\mathbb{E} \left[ \exp \left( \frac{1}{2} X^\top A X \right) \right] = \det(I_d - \sigma^2 A)^{-1/2}.$$

- 8.a. Si  $n = 1$  (une seule observation  $X = X_1$ ), montrer que pour tout  $u, v \in \mathbb{S}_{d-1}$ ,

$$\begin{aligned} \mathbb{E}_0[L_u(X)L_v(X)] &= \mathbb{E}_0 \left[ (1 + \theta)^{-1} \exp \left( \frac{\theta}{2\sigma^2(1 + \theta)} \left[ (u^\top X)^2 + (v^\top X)^2 \right] \right) \right] \\ &= \left( 1 - \theta^2 (u^\top v)^2 \right)^{-1/2} \end{aligned}$$

- 8.b. On note  $L$  le rapport de vraisemblance entre  $\mathbb{P}_1$  et  $\mathbb{P}_0$ . Montrer que  $\chi^2(\mathbb{P}_1, \mathbb{P}_0) = \mathbb{E}_0[L(X)^2] - 1$ , où  $\mathbb{E}_0$  est l'espérance sous  $\mathbb{P}_0$ .

- 8.c. Montrer qu'on a  $L(x) = \mathbb{E}_u[L_u(x)]$ , où  $\mathbb{E}_u$  désigne l'espérance par rapport à  $u \sim \mu$ . En déduire

$$\chi^2(\mathbb{P}_1, \mathbb{P}_0) = \mathbb{E}_{u,v} \left[ \left( 1 - \theta^2 (u^\top v)^2 \right)^{-n/2} \right] - 1,$$

où  $\mathbb{E}_{u,v}$  désigne l'espérance par rapport à  $u, v$  i.i.d. selon  $\mu$ .

9. Dans cette question, toute trace de recherche même partielle, ou réutilisation des résultats précédents, sera prise en compte. On considère le régime où  $n \rightarrow +\infty, d = d(n) \rightarrow +\infty$ . On autorise aussi  $\theta$  à dépendre de  $n$ .

Montrer que si

$$\frac{\theta^2 n}{d} \rightarrow 0,$$

alors le test du problème est infaisable (on définira ce qu'infaisable veut dire).

10. Commenter le résultat établi ci-dessus. On pourra notamment reprendre les conjectures de la question 2.

*Fin du sujet.*